



#5

# SEQUENCE LISTING

<110> Kinsella, Todd

<120> METHODS AND COMPOSITIONS FOR SCREENING USING DIPHTHERIA  
TOXIN CONSTRUCTS

<130> A-70036/RMS/JJD

<140> US 09/712,821

<141> 2000-11-13

<150> US 60/165,189

<151> 1999-11-12

<160> 19

<170> PatentIn Ver. 2.1

<210> 1

<211> 603

<212> DNA

<213> Homo sapiens

<400> 1

```
ctcgaggaca gtgacctggg agtgagtaca aggtgaggcc accactcagg gtgccagctc 60
caagcgggtc acagggacga gggctgcggc catcaggagg ccctgcacac acatctggga 120
cacgcgcccc cgagggccag ttcacctcag tgcgcctcat tctcctgcac aaaagcgccc 180
ccatcctttc ttcacaaggc tttcgtggaa gcagaggcgt cgatgcccag taccctctcc 240
ctttcccagg caacgggacc ccaagtttgc tgactgggac caccaagcca cgcatgcgtc 300
aagagtgaga gtccgggacc taggcagggg ccctgggggt gggcctgaga gagaagagaa 360
cctccccag cactcgggtg gcatcggtag tgaaggagcc tcacctgacc cccgctgttg 420
ctcaatcgac ttccaagaa cagagagaaa agggaacttc cagggcgggc cgggcctcct 480
gggggttccc accccatttt tagctgaaag cactgaggca gagctcccc taccaggct 540
cactgcccg gcacagaaat aacaaccacg gttactgatc atctgggagc tgtccaggaa 600
ttc 603
```

<210> 2

<211> 143

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic

<400> 2

```
gctgggctaa actgggctag cctgagctgg gctgaactgg gctgctgggc tggactgggt 60
```

aagctgggct gagctggggtt ggggtgaaat gggctgagct gagctaggct aaactggggtt 120  
 tggctgggct gggctgggct ggg 143

<210> 3  
 <211> 76  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic

<400> 3  
 ggtttggctg ggctgggctg ggctgggctg ggttcagctg agcgggttgg gttagactgg 60  
 gtcaaactgg ttcagc 76

<210> 4  
 <211> 6219  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: synthetic

<400> 4  
 atcacgaggc cctttcgtct tcaagaacag ctttgctctt aggagtttcc taatacatcc 60  
 caaactcaaa tatataaagc atttgacttg ttctatgccc tagttattaa tagtaatcaa 120  
 ttacggggtc attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 180  
 atggccccgc tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg 240  
 ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggag tatttacggg 300  
 aaactgcca cttggcagta catcaagtgt atcatatgcc aagtagcccc cctattgacg 360  
 tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta tgggactttc 420  
 ctacttgga gtacatctac gtattagtca tcgctattac catggtgatg cggttttggc 480  
 agtacatcaa tgggcgtgga tagcggtttg actcacggg atttccaagt ctccaccca 540  
 ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca aaatgtcgta 600  
 acaactccgc ccattgacg caaatggcg gttagcatgt acggtgggag gtctatataa 660  
 gcagagctca ataaaagagc ccacaacccc tactcgggg cgccagtcct ccgattgact 720  
 gagtcgcccc ggtacccgtg tatccaataa accctcttgc agttgcatcc gacttgtggg 780  
 ctcgctgttc cttgggaggg tctcctctga gtgattgact acccgtcagc ggggggtctt 840  
 catttggggg ctgcgtccgg atcgggagac ccctgcccag ggaccaccga cccaccaccg 900  
 ggaggtgaagc tggccagcaa cttatctgtg tctgtccgat tgtctagtgt ctatgactga 960  
 ttttatgcgc ctgcgtcgg actagttagc taactagctc tgtatctggc ggaccggtgg 1020  
 tggaactgac gagttcgga caccggccg caaccctggg agacgtcca gggacttcgg 1080  
 gggccgtttt tgtggcccga cctgagtcca aaaatcccga tcgttttggc ctctttggtg 1140  
 caccctccct agaggaggga tatgtggttc tggtaggaga cgagaaccta aaacagttcc 1200  
 cgctccgctc tgaatttttg ctttcggttt gggaccgaag ccgcgccg cgctctgtct 1260  
 gctgcagcat cgttctgtgt tgtctctgtc tgactgtgtt tctgtatttg tctgaaaata 1320

tcggccccggg	ccagactgtt	accactccct	taagtttgac	cttaggtcac	tggaaagatg	1380
tcgagcggat	cgctcacaac	cagtcggtag	atgtcaagaa	gagacgttgg	gttaccttct	1440
gctctgcaga	atggccaacc	tttaacgtcg	gatggccgcg	agacggcacc	tttaaccgag	1500
acctcatcac	ccaggttaag	atcaaggtct	tttcacctgg	cccgcatgga	caccagacc	1560
aggtccccta	catcgtgacc	tgggaagcct	tggcttttga	ccccctccc	tgggtcaagc	1620
cctttgtaca	ccctaagcct	ccgcctcctc	ttcctccatc	cgccccgtct	ctcccccttg	1680
aacctcctcg	ttcgaccccc	cctcgatcct	cccttttatcc	agccctcact	ccttctctag	1740
gcgcccccat	atggccatat	gagatcttat	atggggcacc	cccgccccct	gtaaacttcc	1800
ctgaccctga	catgacaaga	gttactaaca	gccccctctc	ccaagctcac	ttacaggctc	1860
tctacttagt	ccagcacgaa	gtctggagac	ctctggcggc	agcctaccaa	gaacaactgg	1920
accgaccggt	ggtacctcac	ccttaccgag	tcggcgacac	agtgtgggtc	cgccgacacc	1980
agactaagaa	cctagaacct	cgctggaaag	gaccttacac	agtcctgctg	accacccccca	2040
ccgcccccaa	agtagacggc	atcgcgcttg	gatacacgcc	gcccacgtga	aggctgccga	2100
ccccgggggt	ggaccatcct	ctagactgcc	ggatctcgag	ggatccacca	ccatggaccc	2160
ccattaaatt	ggaattcctg	cagccccggg	gatccactag	ttctagagcg	aattaattcc	2220
ggttattttc	caccatattg	ccgtcttttg	gcaatgtgag	ggcccggaag	cctggccctg	2280
tcttcttgac	gagcattcct	aggggtcttt	cccctctcgc	caaaggaatg	caaggtctgt	2340
tgaatgtcgt	gaaggaagca	gttcctctgg	aagcttcttg	aagacaaaca	acgtctgtag	2400
cgaccctttg	caggcagcgg	aaccccccac	ctggcgacag	gtgcctctgc	ggccaaaagc	2460
cacgtgtata	agatacacct	gcaaaggcgg	cacaaccca	gtgccacgtt	gtgagttgga	2520
tagttgtgga	aagagtcaaa	tggctctcct	caagcgtatt	caacaagggg	ctgaaggatg	2580
cccagaaggt	accccatgtg	atgggatctg	atctggggcc	tcggtgcaca	tgctttacat	2640
gtgttttagtc	gaggttaaaa	aacgtctagg	ccccccgaac	cacggggacg	tggttttctc	2700
ttgaaaaaca	cgatgataat	atgggggatc	caccggtcgc	caccatgggtg	agcaagggcg	2760
aggagctgtt	caccgggggtg	gtgcccattc	tggctgagct	ggacggcgac	gtaaacggcc	2820
acaagttcag	cgtgtccggc	gagggcgagg	gcgatgccac	ctacggcaag	ctgaccctga	2880
agttcatctg	caccaccggc	aagctgcccg	tgccttgccc	caccctcgtg	accaccctga	2940
cctacggcgt	gcagtgttcc	agccgctacc	ccgaccacat	gaagcagcac	gacttcttca	3000
agtcggccat	gcccgaaggc	tacgtccagg	agcgcaccat	cttcttcaag	gacgacggca	3060
actacaagac	ccgcgcggag	gtgaagttcg	agggcgacac	cctggtgaac	cgcatcgagc	3120
tgaagggcat	cgacttcaag	gaggacggca	acatcctggg	gcacaagctg	gagtacaact	3180
acaacagcca	caacgtctat	atcatggccg	acaagcagaa	gaacggcatc	aaggtgaact	3240
tcaagatccg	ccacaacatc	gaggacggca	gcgtgcagct	cgccgaccac	taccagcaga	3300
acacccccat	cggcgacggc	cccggtgctgc	tgcccagaaa	ccactacctg	agcaccagct	3360
ccgcccctgag	caaagacccc	aacgagaagc	gcgatcacat	ggtcctgctg	gagttcgtga	3420
ccgcccgcgg	gatcactctc	ggcatggacg	agctgtacaa	gtaaagcggc	cgctcgacga	3480
taaaataaaa	gattttattt	agtctccaga	aaaagggggg	aatgaaagac	cccacctgta	3540
ggtttggcaa	gctagcttaa	gtaacgcat	tttgcaaggc	atggaaaaat	acataactga	3600
gaatagagaa	gttcagatca	aggtcaggaa	cagatggaac	agctgaatat	gggccaacaa	3660
ggatatctgt	ggtaagcagt	tcctgccccg	gctcagggcc	aagaacagat	ggaacagctg	3720
aatatgggcc	aaacaggata	tctgtggtaa	gcagttcctg	ccccggctca	gggccaagaa	3780
cagatggtcc	ccagatgcgg	tccagccctc	agcagtttct	agagaacat	cagatgtttc	3840
cagggtgccc	caaggacctg	aaatgacctt	gtgccttatt	tgaactaacc	aatcagttcg	3900
cttctcgctt	ctgttcgcgc	gcttctgctc	cccagactca	ataaaagagc	ccacaacccc	3960
tactcggggg	cgccagtcct	ccgattgact	gagtcgcccg	ggtaccctgt	tatccaataa	4020
accctcttgc	agttgcatcc	gacttgtggt	ctcgtgttcc	cttgggaggg	tctcctctga	4080
gtgattgact	acccgtcgcg	ggggctcttc	atttccgact	tgtgggtctg	ctgccttggg	4140
agggctctct	ctgagtgatt	gactaccctg	cagcgggggt	cttcacatgc	agcatgtatc	4200

```

aaaattaatt tgggtttttt tcttaagtat ttacattaaa tggccatagt tgcattaatg 4260
aatcggccaa cgcgcgggga gaggcgggtt gcgatttggc gctcttccgc ttcctcgctc 4320
actgactcgc tgcgctcggg cggttcggctg cggcgagcgg tatcagctca ctcaaaggcg 4380
gtaatacggg tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 4440
cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc 4500
ccccctgacg agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 4560
ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc 4620
ctgccgctta cgggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat 4680
agctcacgct gtaggtatct cagttcgggt taggtcgttc gctccaagct gggctgtgtg 4740
cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 4800
aaccggtaaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 4860
gcgaggatag taggcgggtg tacagagttc ttgaagtggg ggcctaacta cggctacact 4920
agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 4980
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag 5040
cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 5100
tctgacgctc agtggaaacg aaactcacgt taagggattt tggatcatgag attatcaaaa 5160
aggatcttca cctagatcct tttaaattaa aaatgaagtt tgcgcaaatc aatctaaagt 5220
atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca 5280
gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg 5340
atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca 5400
ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggg 5460
cctgcaactt tatccgcctc catccagtct attaattggt gccgggaagc tagagtaagt 5520
agttcgccag ttaatagttt gcgcaacggt gttgccattg ctacaggcat cgtggtgtca 5580
cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca 5640
tgatccccc tgttggtgaa aaaagcgggt agctccttcg gtctcctgat cgttgctcaga 5700
agtaagtggg ccgcagtggt atcactcatg gttatggcag cactgcataa ttctcttact 5760
gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcatcttga 5820
gaatagtgtg tgcggcgacc gagttgctct tgcccggcgt caacacggga taataccgcg 5880
ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc 5940
tcaaggatct taccgctggt gagatccagt tcgatgtaac ccactcgtgc acccaactga 6000
tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat 6060
gccgcaaaaa agggaataag ggcgacagg aaatgttgaa tactcatact cttccttttt 6120
caatattatt gaagcattta tcaggttatt gtctcatgag cggatacata tttgaatgta 6180
tttagaaaaa taaacaaata ggggttccgc gcacatttc 6219

```

<210> 5

<211> 5713

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic

<400> 5

```

atcacgaggc cctttcgtct tcaagaacag ctttgctctt aggagtttcc taatacatcc 60
caaactcaaa tatataaagc atttgacttg ttctatgcc tagttattaa tagtaataca 120
ttacgggggc attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 180

```

atggccccgc	tggctgaccg	cccaacgacc	cccgcccatt	gacgtcaata	atgacgtatg	240
ttcccatagt	aacgcccaata	gggactttcc	attgacgtca	atgggtggag	tattttacggt	300
aaactgcccc	cttggcagta	catcaagtgt	atcatatgcc	aagtacgccc	cctattgacg	360
tcaatgacgg	taaatggccc	gcctggcatt	atgcccagta	catgacctta	tgggactttc	420
ctacttgga	gtacatctac	gtattagtca	tcgctattac	catggtgatg	cggttttggc	480
agtacatcaa	tgggcgtgga	tagcggtttg	actcacgggg	atttccaagt	ctccacccca	540
ttgacgtcaa	tgggagtttg	ttttggcacc	aaaatcaacg	ggactttcca	aaatgtcgta	600
acaactccgc	cccattgacg	caaatgggcg	gtaggcatgt	acggtgggag	gtctatataa	660
gcagagctca	ataaaagagc	ccacaacccc	tcactcgggg	cgccagtcct	ccgattgact	720
gagtcgcccc	ggtaccctgt	tatccaataa	accctcttgc	agttgcatcc	gacttgtggt	780
ctcgtgttcc	cttgggaggg	tctcctctga	gtgattgact	acccgtcagc	gggggtcttt	840
catttggggg	ctcgtccggg	atcgggagac	ccctgcccag	ggaccaccga	cccaccaccg	900
ggaggtgaag	tggccagcaa	cttatctgtg	tctgtccgat	tgtctagtgt	ctatgactga	960
ttttatgcgc	ctgcgtcggg	actagttagc	taactagctc	tgtatctggc	ggaccctgtg	1020
tggaactgac	gagttcgga	caccggcgcc	caaccctggg	agacgtccca	gggacttcgg	1080
gggccgtttt	tgtggcccga	cctgagtcca	aaaatcccga	tcgttttggg	ctctttggtg	1140
cacccccctt	agaggaggga	tatgtggttc	tggtaggaga	cgagaacctt	aaacagttcc	1200
cgcctccgtc	tgaatttttg	ctttcggttt	gggaccgaag	ccgcgcgcgc	cgtcttgtct	1260
gctgcagcat	cgttctgtgt	tgtctctgtc	tgactgtgtt	tctgtatttg	tctgaaaata	1320
tcggccccgg	ccagactgtt	accactccct	taagtttgac	cttaggtcac	tggaaagatg	1380
tcgagcggat	cgctcacaac	cagtcggtag	atgtcaagaa	gagacgttgg	gttaccttct	1440
gctctgcaga	atggccaacc	tttaacgtcg	gatggccgcg	agacggcacc	tttaaccgag	1500
acctcatcac	ccaggttaag	atcaaggtct	tttcacctgg	cccgcattga	caccagacc	1560
aggtccccta	catcgtgacc	tgggaagcct	tggcttttga	ccccctccc	tgggtcaagc	1620
cctttgtaca	ccctaagcct	ccgcctcctc	ttcctccatc	cgccccgtct	ctcccccttg	1680
aacctcctcg	ttcgaccccc	cctcgatcct	ccctttatcc	agccctcact	ccttctctag	1740
gcgcccccat	atggccatat	gagatcttat	atggggcacc	cccgccccct	gtaaacttcc	1800
ctgaccctga	catgacaaga	gttactaaca	gccccctctc	ccaagctcac	ttacaggctc	1860
tctacttagt	ccagcacgaa	gtctggagac	ctctggcggc	agcctacca	gaacaactgg	1920
accgaccggt	ggtacctcac	ccttaccgag	tcggcgacac	agtgtgggtc	cgccgacacc	1980
agactaagaa	cctagaacct	cgctggaaag	gaccttacac	agtcctgctg	accaccccca	2040
ccgcccctca	gtagacggca	tcgcagcttg	gatacacgcc	gccacagtga	aggctgccga	2100
ccccgggggt	ggaccatcct	ctagactgcc	ggatctcgag	ggatccacca	tggtagacaa	2160
gggcgaggag	ctgttcaccg	gggtggtgcc	catcctggtc	gagctggacg	gcgacgtaaa	2220
cggccacaag	ttcagcgtgt	ccggcgaggg	cgagggcgat	gccacctacg	gcaagctgac	2280
cctgaagttc	atctgcacca	ccggcaagct	gcccgtgccc	tggcccaccc	tcgtgaccac	2340
cctgacctac	ggcgtgcagt	gcttcagccg	ctaccccgac	cacatgaagc	agcacgactt	2400
cttcaagtcc	gccatgcccc	aaggctacgt	ccaggagcgc	accatcttct	tcaaggacga	2460
cggcaactac	aagaccgcgc	ccgaggtgaa	gttcgagggc	gacaccctgg	tgaaccgcat	2520
cgagctgaag	ggcatcgact	tcaaggagga	cggcaacatc	ctggggcaca	agctggagta	2580
caactacaac	agccacaacg	tctatatcat	ggccgacaag	cagaagaacg	gcataaggt	2640
gaacttcaag	atccgccaca	acatcgagga	cggcagcgtg	cagctcgccg	accactacca	2700
gcagaacacc	cccatcggcg	acggccccgt	gctgctgccc	gacaaccact	acctgagcac	2760
ccagtccgcc	ctgagcaaag	accccaacga	gaagcgcgat	cacatggtcc	tgctggagtt	2820
cgtgaccgcc	gccgggatca	ctctcggcgt	ggacgagctg	tacaaggaat	tcggaggtgg	2880
cagcgggtgg	ggtcagctgt	tgaattttga	ccttcttaaa	cttgcgggag	acgtcgagtc	2940
caaccctggg	cccaccacca	ccatggaagc	ttccattaaa	ttggttaacg	tcgacgcggc	3000
cgctcgacga	taaaataaaa	gattttatatt	agtctccaga	aaaagggggg	aatgaaagac	3060

cccacctgta	ggtttggcaa	gctagcttaa	gtaacgccat	tttgcaaggc	atggaaaaat	3120
acataactga	gaatagagaa	gttcagatca	aggtcaggaa	cagatggaac	agctgaatat	3180
gggccaacaa	ggatatctgt	ggtaagcagt	tcctgccccg	gctcagggcc	aagaacagat	3240
ggaacagctg	aatatgggcc	aaacaggata	tctgtggtaa	gcagttcctg	ccccggctca	3300
gggccaagaa	cagatgggtcc	ccagatgcgg	tccagccctc	agcagtttct	agagaaccat	3360
cagatgtttc	caggggtgccc	caaggacctg	aaatgaccct	gtgccttatt	tgaactaacc	3420
aatcagttcg	cttctcgctt	ctgttcgcgc	gcttctgctc	cccagactca	ataaaagagc	3480
ccacaacccc	tactcggggg	cgccagtcct	ccgattgact	gagtcgcccc	ggtagccgtg	3540
tatccaataa	accctcttgc	agttgcatcc	gacttggtgt	ctcgctgttc	cttgggaggg	3600
tctcctctga	gtgattgact	accgcgcagc	gggggtcttt	catttccgac	ttgtggtctc	3660
gctgccttgg	gaggggtctcc	tctgagtgat	tgactaccgc	tcagcggggg	tcttcacatg	3720
cagcatgtat	caaaattaat	ttggtttttt	ttcttaagta	tttacattaa	atggccatag	3780
ttgcattaat	gaatcgcca	acgcgcgggg	agaggcggtt	tgcgatttgg	cgctcttccg	3840
cttcctcgct	cactgactcg	ctgcgctcgg	tcgttcggct	gcggcgagcg	gtatcagctc	3900
actcaaaggc	ggtaatacgg	ttatccacag	aatcagggga	taacgcagga	aagaacatgt	3960
gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	cgcgttgctg	gcgtttttcc	4020
ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	ctcaagtcag	aggtggcgaa	4080
acccgacagg	actataagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	4140
tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	4200
gctttctcat	agctcacgct	gtaggtatct	cagttcgggt	taggtcgttc	gctccaagct	4260
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	4320
tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	4380
gattagcaga	gcgaggtatg	taggcggtgc	tacagagttc	ttgaagtggg	ggcctaacta	4440
cggctacact	agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	4500
aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	4560
tgtttgcaag	cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	4620
ttctacgggg	tctgacgctc	agtggaacga	aaactcacgt	taagggtatt	tggtcatgag	4680
attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	aaatgaagtt	tgcgcaaata	4740
aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagtgaggc	4800
acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcc	ccgtcgtgta	4860
gataactacg	atacgggagg	gcttaccatc	tggccccagt	gctgcaatga	taccgcgaga	4920
cccacgctca	ccggctccag	atttatcagc	aataaaccag	ccagccggaa	gggccgagcg	4980
cagaagtggg	cctgcaactt	tatccgcctc	catccagttc	attaattggt	gccgggaagc	5040
tagagtaagt	agttcgccag	ttaatagttt	gcgcaacggt	gttgccattg	ctacaggcat	5100
cgtgggtgtca	cgctcgtcgt	ttggtatggc	ttcattcagc	tccggttccc	aacgatcaag	5160
gcgagttaca	tgatcccca	tgttgtgcaa	aaaagcggtt	agctccttcg	gtcctccgat	5220
cgttgtcaga	agtaagttgg	ccgcagtgtt	atcactcatg	gttatggcag	cactgcataa	5280
ttctcttact	gtcatgccat	ccgtaagatg	cttttctgtg	actggtgagt	actcaaccaa	5340
gtcattctga	gaatagtgtg	tgccggcgacc	gagttgctct	tgcccgccgt	caacacggga	5400
taataccgcg	ccacatagca	gaactttaaa	agtgtctatc	attggaaaac	gttcttcggg	5460
gcgaaaactc	tcaaggatct	taccgctggt	gagatccagt	tcgatgtaac	ccactcgtgc	5520
acccaactga	tcttcagcat	cttttacttt	caccagcggt	tctgggtgag	caaaaacagg	5580
aaggcaaaat	gccgcaaaaa	aggggaataag	ggcgacacgg	aaatgttgaa	tactcatact	5640
cttccttttt	caatattatt	gaagcattta	tcaggggtat	tgtctcatga	cattaacctt	5700
taaaaatagg	cgt					5713

<211> 4922

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic

<400> 6

```
atcacgaggc cctttcgtct tcaagaacag ctttgctctt aggagtttcc taatacatcc 60
caaactcaaa tatataaagc atttgacttg ttctatgcc tagttattaa tagtaatcaa 120
ttacgggggc attagttcat agccatatat ggagttccgc gttacataac ttacggtaaa 180
tggcccgcc ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt 240
tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta 300
aactgcccac ttggcagtac atcaagtgt tcatatgcc agtacgccc ctattgacgt 360
caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc 420
tacttggcag tacatctacg tattagtcac cgctattacc atggtgatgc ggttttggca 480
gtacatcaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat 540
tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa 600
caactccgcc ccattgacgc aaatgggagg taggcattga cgggtggagg tctatataag 660
cagagctcaa taaaagagcc cacaaccctt cactcggggc gccagtctc cgattgactg 720
agtcgcccgg gtaccggtgt atccaataaa ccctcttgca gttgcatccg acttgtggtc 780
tcgctgttcc ttgggagggg ctccctctgag tgattgacta cccgtcagcg ggggtcttcc 840
atttgggggc tcgtccggga tcgggagacc cctgccagg gaccaccgac ccaccaccgg 900
gaggtaaagt ggccagcaac ttatctgtgt ctgtccgatt gtctagtgtc tatgactgat 960
tttatgcgcc tgcgtcggtg ctagttagct aactagctct gtatctggcg gaccctgggt 1020
ggaactgacg agttcggaac acccgggcgc aaccctggga gacgtcccag ggacttcggg 1080
ggccgttttt gtggcccgac ctgagtccaa aaatcccgat cgttttggac tctttgggtg 1140
acccccctta gaggagggat atgtggttct ggtaggagac gagaacctaa aacagttccc 1200
gcctccgtct gaatttttgc tttcggtttg ggaccgaagc cgcgccgcgc gtcttgtctg 1260
ctgcagcatc gttctgtgtt gtctctgtct gactgtgttt ctgtatttgt ctgaaaatat 1320
cggcccgggc cagactgtta ccaactccct aagtttgacc ttaggtcact ggaaagatgt 1380
cgagcggatc gtcacaacc agtcggtaga tgtcaagaag agacgttggg ttaccttctg 1440
ctctgcagaa tggccaacct ttaacgtcgg atggccgca gacggcacct ttaaccgaga 1500
cctcatcacc caggttaaga tcaaggctct ttacactggc ccgcatggac acccagacca 1560
ggtcccctac atcgtgacct gggaagcctt ggcttttgac cccctccctt ggggtcaagc 1620
ctttgtacac cctaagcctc cgctcctctt tctccatcc gcccgtctc tcccccttga 1680
acctcctcgt tcgaccccgcc ctcgatcctc cctttatcca gccctcactc cttctctagg 1740
cgcccccata tggccatatg agatcttata tggggcaccc ccgccccttg taaacttccc 1800
tgaccctgac atgacaagag ttactaacag cccctctctc caagctcact tacaggctct 1860
ctacttagtc cagcacgaag tctggagacc tctggcgga gcctaccaag aacaactgga 1920
ccgaccggtg gtacctcacc cttaccgagt cggcgacaca gtgtgggtcc gccgacacca 1980
gactaagaac ctagaacctc gctggaaagg accttacaca gtctgtctga ccacccccac 2040
cgccctcaag tagacggcat cgcagcttgg atacacgccg cccacgtgaa ggctgccgac 2100
cccgggggtg gaccatcctc tagactgccg gatctcgagg gatccaccac catggacccc 2160
cattaaattg gaattcgggg cccaagcttt gttaacgtcg acgcggccgc cgtcgacgat 2220
aaaataaaag attttattta gtctccagaa aaagggggga atgaaagacc ccacctgtag 2280
gtttggcaag ctagcttaag taacgccatt ttgcaaggca tggaaaaata cataactgag 2340
aatagagaag ttcagatcaa ggtcaggaac agatggaaca gctgaatatg ggccaaacag 2400
```

gatatctgtg	gtaagcagtt	cctgccccgg	ctcagggcca	agaacagatg	gaacagctga	2460
atatgggcca	aacaggatat	ctgtggtaag	cagttcctgc	cccggctcag	ggccaagaac	2520
agatgggtccc	cagatgcggt	ccagccctca	gcagtttcta	gagaaccatc	agatgtttcc	2580
agggtgcccc	aaggacctga	aatgacctg	tgccttattt	gaactaacca	atcagttcgc	2640
ttctcgcttc	tgttcgcgcg	cttctgctcc	ccgagctcaa	taaaagagcc	cacaaccct	2700
cactcggggc	gccagtcctc	cgattgactg	agtcgccccg	gtaccctgtg	atccaataaa	2760
ccctcttgca	gttgcatccg	acttggtggtc	tcgctgttcc	ttgggagggg	ctcctctgag	2820
tgattgacta	cccgtcagcg	ggggtctttc	atttcagact	tgtggtctcg	ctgccttggg	2880
agggtctcct	ctgagtgatt	gactaccctg	cagcgggggt	cttcacatgc	agcatgtatc	2940
aaa'attaatt	tggttttttt	tcttaagtat	ttacattaaa	tggccatagt	tgcattaatg	3000
aatcggccaa	cgcgcgggga	gaggcggttt	gcgtattggc	gctcttccgc	ttcctcgctc	3060
actgactcgc	tgcgctcggg	cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	3120
gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaaggc	3180
cagcaaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	3240
ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	gggtggcgaaa	cccgcacagga	3300
ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	tgttcgcgacc	3360
ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	3420
agctcacgct	gtaggtatct	cagttcgggtg	taggtcggtc	gctccaagct	gggctgtgtg	3480
cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	3540
aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	3600
gcgaggtatg	taggcgggtgc	tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	3660
agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	3720
ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	tgtttgcaag	3780
cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	3840
tctgacgctc	agtggaaacga	aaactcacgt	taagggattt	tgggtcatgag	attatcaaaa	3900
aggatcttca	cctagatcct	tttaaattaa	aaatgaagtt	tgcgcaaatc	aatctaaagt	3960
atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagtgaggc	acctatctca	4020
gcgatctgtc	tatttcggtc	atccatagtt	gcctgactcc	ccgtcgtgta	gataactacg	4080
atacgggagg	gcttacatct	ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	4140
cggctccaga	tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	agaagtgggtc	4200
ctgcaacttt	atccgcctcc	atccagtcta	ttaattggtg	ccgggaagct	agagtaagta	4260
gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	tacaggcatc	gtggtgtcac	4320
gctcgtcggt	tggatatggc	tcattcagct	ccggttccca	acgatcaagg	cgagttacat	4380
gatcccccat	gttgtgcaaa	aaagcgggtta	gctccttcgg	tcctccgata	gttgtcagaa	4440
gtaagtgggc	cgcagtggtta	tcaactcatg	ttatggcagc	actgcataat	tctcttactg	4500
tcatgccatc	cgtaagatgc	ttttctgtga	ctggtgagta	ctcaaccaag	tcattctgag	4560
aatagtgtat	gcggcgaccg	agttgctctt	gccccgcgtc	aacacgggat	aataccgcgc	4620
cacatagcag	aactttaaaa	gtgctcatca	ttggaaaacg	ttcttcgggg	cgaaaactct	4680
caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	cactcgtgca	cccaactgat	4740
cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	aaaaacagga	aggcaaaatg	4800
ccgcaaaaaa	gggaataagg	gcgacacgga	aatgttgaat	actcatactc	ttcctttttc	4860
aatattattg	aagcatttat	cagggttatt	gtctcatgac	attaacctat	aaaaataggc	4920
gt						4922

<210> 7

<211> 8282

<212> DNA



<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic

<400> 7

```
atcacgaggc cctttcgtct tcaagaacag ctttgctctt aggagtttcc taatacatcc 60
caaactcaaa tatataaagc atttgacttg ttctatgccc tagttattaa tagtaatcaa 120
ttacggggtc attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 180
atggcccgcg tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg 240
ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggag tatttacggg 300
aaactgcccc cttggcagta catcaagtgt atcatatgcc aagtacgccc cctattgacg 360
tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta tgggactttc 420
ctacttgga gtacatctac gtattagtca tcgctattac catggtgatg cggttttggc 480
agtacatcaa tgggcgtgga tagcggtttg actcacggg atttccaagt ctccaccca 540
ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca aaatgtcgta 600
acaactccgc ccctattgacg caaatgggcg gtaggcattg acggtgggag gtctatataa 660
gcagagctca ataaaagagc ccacaacccc tactcgggg cgccagtcct ccgattgact 720
gagtcgcccg ggtaccctg tatccaataa accctcttgc agttgcatcc gacttgtggg 780
ctcgtgttc cttgggaggg tctcctctga gtgattgact acccgtcagc gggggctctt 840
catttggggg ctcgtccggg atcgggagac ccctgccag ggaccaccga cccaccaccg 900
ggaggtaagc tggccagcaa cttatctgtg tctgtccgat tgtctagtgt ctatgactga 960
ttttatgcgc ctgcgtcggg actagttagc taactagctc tgtatctggc ggaccctggg 1020
tggaactgac gagttcggaa caccggccg caaccctggg agacgtcca gggacttcgg 1080
gggccgtttt tgtggcccga cctgagtcca aaaatcccga tcgttttga ctctttggg 1140
caccctctt agaggaggga tatgtggttc tggtaggaga cgagaacctt aaacagttcc 1200
cgctccgctc tgaatttttg ctttcggttt gggaccgaag ccgcgccgcg cgtcttgtct 1260
gctgcagcat cgttctgtgt tgtctctgtc tgactgtgtt tctgtatttg tctgaaaata 1320
tgggcccggg ccagactgtt accactccct taagtttgac cttaggtcac tggaaagatg 1380
tcgagcggat cgctcacaac cagtcggtag atgtcaagaa gagacgttgg gttaccttct 1440
gctctgcaga atggccaacc tttaacgtcg gatggccgcg agacggcacc tttaaccgag 1500
acctcatcac ccaggttaag atcaaggtct tttcacctgg cccgcatgga caccagacc 1560
aggtccccta catcgtgacc tgggaagcct tggcttttga cccctctccc tgggtcaagc 1620
cctttgtaca ccctaagcct ccgcctctc ttcctccatc cgcctcgtct ctcccccttg 1680
aacctcctcg ttcgaccccg cctcgatcct ccctttatcc agccctcact ccttctctag 1740
gcgcccccat atggccatat gagatcttat atggggcacc cccgcccctt gtaaaccttc 1800
ctgaccctga catgacaaga gttactaaca gcccctctct ccaagctcac ttacaggctc 1860
tctacttagt ccagcacgaa gtctggagac ctctggcggc agcctaccaa gaacaactgg 1920
accgaccggt ggtacctcac ccttaccgag tcggcgacac agtgtgggtc cgccgacacc 1980
agactaagaa cctagaacct cgctggaaag gaccttacac agtcctgtg accaccccca 2040
ccgcccctca agtagacggc atcgcagctt ggatacacgc cgcccacgtg aaggctgccg 2100
accccggggg tggaccatcc tctagactgc cggatctcga gggatcctcc ccagcatgcc 2160
tgctattgtc ttccaatcc tcccccttgc tgtcctgccc caccacacc cccagaatag 2220
aatgacacct actcagacaa tgcgatgcaa tttcctcatt ttattaggaa aggacagtgg 2280
gagtggcacc ttccagggtc aaggaaggca cgggggaggg gcaaacaaca gatggctggc 2340
aactagaagg cacagtcgag gtctagcttg ccaaacctac aggtggggtc tttcattccc 2400
ccctttttct ggagactaaa taaaatcttt tattttatcg atagatccc gtcggcatct 2460
actctattcc tttgccctcg gacgagtgt ggggcgtcgg tttccactat cggcgagtac 2520
```

ttctacacag ccatcggtcc agacggccgc gcttctgctg gcgatttgtg tacgcccagc 2580  
 agtcccggct ccggatcgga cgattgctgc gcatcgaccc tgcgcccagg ctgcatcatc 2640  
 gaaattgccg tcaaccaagc tctgatagag ttggtcaaga ccaatgcgga gcatatacgc 2700  
 ccggagccgc ggcgatcctg caagctccgg atgcctccgc tcgaagtagc gcgtctgctg 2760  
 ctccatacaa gccaaaccag gcctccagaa gaagatgttg gcgacctcgt attgggaatc 2820  
 cccgaacatc gcctcgctcc agtcaatgac cgctgttatg cggccattgt ccgtcaggac 2880  
 attgttggag ccgaaatccg cgtgcacgag gtgcgggact tcggggcagt cctcggccca 2940  
 aagcatcagc tcatcgagag cctgcgcgac ggacgcactg acggtgtcgt ccatcacagt 3000  
 ttgccagtga tacacatggg gatcagcaat cgcgcataatg aaatcacgcc atgtagtgtg 3060  
 ttgaccgatt ccttgcggtc cgaatgggcc gaaccgcgtc gtctggctaa gatcggccgc 3120  
 agcgatcgca tccatggcct ccgcgaccgg ctgcagaaca gcgggcagtt cggtttcagg 3180  
 caggtcttgc aacgtgacac cctgtgcacg gcgggagatg caataggtca ggctctcgtc 3240  
 aaattcccca atgtcaagca cttccggaat cgggagcgcg gccgatgcaa agtgccgata 3300  
 aacataacga tctttgtaga aaccatcggc gcagctatct acccgagga catatccacg 3360  
 ccctcctaca tcgaagctga aagcacgaga ttcttcgccc tccgagagct gcatcaggtc 3420  
 ggagacgctg tcgaactttt cgatcagaaa cttctcgaca gacgtcgcgg tgagttcagg 3480  
 ctttttcatg gtattatcat cgtgtttttc aaaggaaaac cagtcctccg tggttcgggg 3540  
 ggcctagacg ttttttaacc tcgactaaac acatgtaaag catgtgcacc gagggcccag 3600  
 atcagatccc atacaatggg gtaccttctg ggcatccttc agccccttgt tgaatacgtc 3660  
 tgaggagagc catttgactc tttccacaac tatccaactc acaacgtggc actgggggtg 3720  
 tgccgccttt gcaggtgtat cttatacacg tggcttttgg ccgcagaggc acctgtcgcc 3780  
 aggtgggggg ttccgctgcc tgcaaagggc cgctacagac gttgtttgtc ttcaagaagc 3840  
 ttccagagga actgcttcct tcacgacatt caacagacct tgcattcctt tggcgagagg 3900  
 ggaaagaccc ctgactaga ccaagctttg gatttcattt ctgaagtttg aattttctga 3960  
 gtcactagta atgtccttga ggatgatagt ctgaattttc tctgcaagag tacaagatt 4020  
 ggcttttttg agatctttaa tcaatgtgtc atacgcttct ttctttccat gaagttgatg 4080  
 ccaattacga agcagttgaa ctttctgttc tgctgtgtct tggacattgt cattcttgat 4140  
 ctcatctatt ttggcttcat tgacaccatt ctttcgaaca aagcctttaa cttgacttag 4200  
 tgtcatgact ccagcaatag tggatgataa tttactcaag tcaacatcag ataaatttat 4260  
 tgccactgtt tcaggattta aggttggaga ttcatagaga ccttggtttt ctttctgtg 4320  
 ctttctgcat gttttctgta cttcctttct cttcacccaa acaattagtg gaattggcaa 4380  
 aagaagaaga caaagccacc ccaaccggtt tctgggactt tgtttcctgc agtttgtatt 4440  
 gctggttgct gtgcatggct caagggttcc atgttcacac gaggcgcagc gaacacagtg 4500  
 ttcacagcca ggagaatcgc agtagaagtc tggtttgcac ttgcacttgg tattctgggt 4560  
 caggggtgcag tttgtttcca cttctaaacc atgctcttca tcgcagagtg tgcatcttct 4620  
 gcatttatca gcataatggg tcttgtccat gtactccttc ccttctgtgc atggggcaca 4680  
 ggttgggtga ccccatccta ttttgcagtc ctcaactttt tttttaccag gttggcatgg 4740  
 ttgacagcaa aatgggcctc cttgatataa tccttctgag cagtttttat cagtttcatg 4800  
 aaccgcctc ctcagcttta aactctcgga gatgctatta gtaccttgag tatgaactct 4860  
 taactgtgag ccagcaagca ccagaggcag gacagcccag atccacacca tgggtggcttt 4920  
 accaacagta ccggaatgcc aagcttgcgg ccgcttaaga gctgtaattg aacctgggag 4980  
 tggacacctg tggagagaaa ggcaaagtgg atgtcagtaa gaccaatagg tgcctatcag 5040  
 aaacgcaaga gtcttctctg tctcgacaag ccagtttctt attggtctcc ttaaacctgt 5100  
 cttgtaacct tgatacttac ctgcccagtg cctcacgacc aacttctgca ggaattcctg 5160  
 gacagctccc agatgatcag taaccgtggg tgttatttct gtgccgggca gtggagcctg 5220  
 ggtaggggga gctctgcctc agtgctttca gctaaaaatg ggggtgggaac cccagggagg 5280  
 cccgggccgc cctggaagtt cccttttctc tctgttcttg ggaagtcatg tgagcaacag 5340  
 cgggggctcag gtgaggctcc ttcactaccg atgcacaccg agtgctgggg gaggttctct 5400

tctctctcag	gcccaccccc	agggcccctg	cctaggtccc	ggactctcac	tcttgacgca	5460
tgcgtggcct	ggtgggtccca	gtcagcaaac	ttgggggtccc	gttgccctggg	aaagggagag	5520
ggtactgggc	atcgacgcct	ctgcttccac	gaaagccttg	tgaagaaagg	atgggggagc	5580
ttttgtgcag	gagaatgagg	cgactgagg	tgaactggcc	ctcgggggag	cgtgtcccag	5640
atgtgtgtgc	agggcctcct	gatggccgca	gccctcgtcc	ctgtgacccg	cttgagagctg	5700
gcaccctgag	tgggtggcctc	accttggtact	cactcccagg	tcactgtcct	cgacgcggcc	5760
gctcgacgat	aaaataaaaag	attttatttta	gtctccagaa	aaagggggga	atgaaagacc	5820
ccacctgtag	gtttggcaag	ctagcttaag	taaccatttt	tgcaaggcat	ggaaaaatac	5880
ataactgaga	atagagaagt	tcagatcaag	gtcggaaacag	atggaacagg	caataaaaga	5940
gccacaacc	cctcactcgg	ggcgccagtc	ctccgattga	ctgagtcgcc	cgggtacccg	6000
tgtatccaat	aaaccctctt	gcagttgcat	ccgacttgtg	gtctcgtgtg	tccttgggag	6060
ggtctcctct	gagtgtattga	ctaccgcgtca	gcgggggtct	ttcacatgca	gcatgtatca	6120
aaattaattt	ggtttttttt	cttaagtatt	tacattaaat	ggccatagtt	tcgtaatcat	6180
ggtcatagct	gtttcctgtg	tgaaattggt	atccgctcac	aattccacac	aacatacgag	6240
ccggaagcat	aaagtgtaaa	gcctgggggtg	cctaattgagt	gagctaactc	acattaattg	6300
cgttgcgctc	actgcccgct	ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	6360
tcggccaacg	cgcggggaga	ggcggtttgc	gtattgggag	ctcttcgctc	tcctcgtctca	6420
ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcggt	atcagctcac	tcaaaggcgg	6480
taatacgggt	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	6540
agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgctggc	gtttttccat	aggctccgcc	6600
ccctgacga	gcatcacaaa	aatcgacgct	caagtacagag	gtggcgaaac	ccgacaggac	6660
tataaagata	ccaggcggtt	ccccctggaa	gctccctcgt	gcgctctcct	gttccgaccc	6720
tgccgcttac	cggataacctg	tccgcctttc	tcccttcggg	aagcgtggcg	ctttctcata	6780
gctcacgctg	taggtatctc	agttcggtgt	aggctcgttcg	ctccaagctg	ggctgtgtgc	6840
acgaaccccc	cgttcagccc	gaccgctgcg	ccttatccgg	taactatcgt	cttgagtcca	6900
accggttaag	acacgactta	tcgccactgg	cagcagccac	tggtaacagg	attagcagag	6960
cgaggtagtg	aggcggtgct	acagagttct	tgaagtgggtg	gcctaactac	ggctacacta	7020
gaaggacagt	atttggtatc	tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	7080
gtagctcttg	atccggcaaa	caaaccaccg	ctggtagcgg	tggttttttt	gtttgcaagc	7140
agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgatcttt	tctacggggt	7200
ctgacgctca	gtggaacgaa	aactcacggt	aagggtatgt	ggtcatgaga	ttatcaaaaa	7260
ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	gcgcaaatca	atctaaagta	7320
tatatgagta	aacttgggtc	gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	7380
cgatctgtct	atctcgttca	tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	7440
tacgggaggg	cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	7500
cggctccaga	tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	agaagtggtc	7560
ctgcaacttt	atccgcctcc	atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	7620
gttcgccagt	taatagtttg	cgcaacgttg	ttgccattgc	tacaggcatc	gtggtgtcac	7680
gctcgtcgtt	tggtagtggt	tcattcagct	ccggttccca	acgatcaagg	cgagttacat	7740
gatcccccat	gttgtgcaaa	aaagcggtta	gctccttcgg	tcctccgata	gttgtcagaa	7800
gtaagtgggc	cgcagtgtta	tcactcatgg	ttatggcagc	actgcataat	tctcttactg	7860
tcatgccatc	cgtaagatgc	ttttctgtga	ctggtgagta	ctcaaccaag	tcattctgag	7920
aatagtgtat	gcggcgaccg	agttgctctt	gcccggcgtc	aacacgggat	aataccgcgc	7980
cacatagcag	aactttaaaa	gtgctcatca	ttggaaaacg	ttcttcgggg	cgaaaactct	8040
caaggatctt	accgctgttg	agatccagtt	cgatgtaacc	cactcgtgca	cccaactgat	8100
cttcagcatc	ttttactttc	accagcgttt	ctgggtgagc	aaaaacagga	aggcaaaatg	8160
ccgcaaaaaa	gggaataaag	gcgacacgga	aatgttgaat	actcatactc	ttcctttttc	8220
aatattattg	aagcatttat	caggggttatt	gtctcatgac	attaacctat	aaaaataggc	8280

&lt;210&gt; 8

&lt;211&gt; 8345

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: synthetic

&lt;400&gt; 8

```

atcacgaggc cctttcgtct tcaagaacag ctttgctctt aggagtttcc taatacatcc 60
caaactcaaa tatataaagc atttgacttg ttctatgccc tagttattaa tagtaatcaa 120
ttacgggggc attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 180
atggcccgcc tggctgaccg cccaacgacc ccgcccatt gacgtcaata atgacgtatg 240
ttcccatagt aacgccataa gggactttcc attgacgtca atgggtggag tatttacggg 300
aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc cctattgacg 360
tcaatgacgg taaatggccc gcctggcatt atgccagta catgacctta tgggactttc 420
ctacttgga gtacatctac gtattagtca tcgctattac catggtgatg cggttttggc 480
agtacatcaa tgggcgtgga tagcggtttg actcacggg atttccaagt ctccacccca 540
ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca aaatgtcgta 600
acaactccgc ccattgacg caaatgggcg gtaggcatt acggtgggag gtctatataa 660
gcagagctca ataaaagagc ccacaacccc tctctgggg cgccagtcct ccgattgact 720
gagtcgcccg ggtaccctg tatccaataa accctcttgc agttgcatcc gacttgtggg 780
ctcgtgttcc cttgggaggg tctcctctga gtgattgact acccgtcagc gggggctctt 840
catttggggg ctcgccggg atcgggagac ccctgccag ggaccaccga cccaccaccg 900
ggaggtaacg tggccagcaa cttatctgtg tctgtccgat tgtctagtgt ctatgactga 960
ttttatgcgc ctgctcggt actagttagc taactagctc tgtatctggc ggaccggtg 1020
tggaactgac gagttcgga caccggccg caaccctggg agacgtccca gggacttcg 1080
gggcccgttt tgtggcccga cctgagtcca aaaatcccga tcgttttgga ctctttggtg 1140
caccctctt agaggaggga tatgtggttc tggtaggaga cgagaacctt aaacagttcc 1200
cgctccgtc tgaatttttg ctttcggttt gggaccgaag ccgcgccgcg cgtcttgtct 1260
gctgcagcat cgttctgtgt tgtctctgtc tgactgtgtt tctgtatttg tctgaaaata 1320
tgggcccggg ccagactgtt accactccct taagtttgac cttaggtcac tggaaagatg 1380
tcgagcggat cgctcacaac cagtcggtag atgtcaagaa gagacgttgg gttaccttct 1440
gctctgcaga atggccaacc tttaacgtcg gatggccgcg agacggcacc tttaaccgag 1500
acctcatcac ccaggttaag atcaaggtct tttcacctgg cccgcatgga caccagacc 1560
aggtccccta catcgtgacc tgggaagcct tggcttttga ccccccctcc tgggtcaagc 1620
cctttgtaca ccctaagcct ccgcctctc ttcctccatc cgccccgtct ctcccccttg 1680
aacctcctcg ttcgaccccg cctcgatcct ccctttatcc agccctcact ccttctctag 1740
gcgcccccat atggccatat gagatcttat atggggcacc cccgcccctt gtaaacttcc 1800
ctgaccctga catgacaaga gttactaaca gcccctctct ccaagctcac ttacaggctc 1860
tctacttagt ccagcacgaa gtctggagac ctctggcggc agcctaccaa gaacaactgg 1920
accgaccggt ggtacctcac cttaccgag tcggcgacac agtgtgggtc cgccgacacc 1980
agactaagaa cctagaacct cgctggaaag gaccttacac agtcctgctg accaccccca 2040
ccgccctcaa agtagacggc atcgcagctt ggatacacgc cgccacgtg aaggctgccg 2100
acccgggggg tggaccatcc tctagactgc cggatctcga gggatcctcc ccagcatgcc 2160

```

tgctattgtc	ttcccaatcc	tcccccttgc	tgtcctgccc	caccccaccc	cccagaatag	2220
aatgacacct	actcagacaa	tgcgatgcaa	tttcctcatt	ttattaggaa	aggacagtgg	2280
gagtggcacc	ttccagggtc	aaggaaggca	cgggggaggg	gcaaacaaca	gatggctggc	2340
aactagaagg	cacagtcgag	gtctagcttg	ccaaacctac	aggtggggtc	tttcattccc	2400
ccctttttct	ggagactaaa	taaaatcttt	tattttatcg	atagatcccg	gtcggcatct	2460
actctattcc	tttgccctcg	gacgagtgtt	ggggcgctcg	tttccactat	cggcgagtac	2520
ttctacacag	ccatcggtcc	agacggccgc	gcttctgctg	gcgatttgtg	tacgcccagc	2580
agtcccggct	ccggatcgga	cgattgctgc	gcacgcaccc	tgcgccaag	ctgcatcatc	2640
gaaattgccg	tcaaccaagc	tctgatagag	ttggtcaaga	ccaatgcgga	gcataacgc	2700
ccggagccgc	ggcgatcctg	caagctccgg	atgcctccgc	tcgaagtagc	gcgtctgctg	2760
ctccatacaa	gccaaccacg	gcctccagaa	gaagatgttg	gcgacctcgt	attgggaatc	2820
cccgaacatc	gcctcgctcc	agtcaatgac	cgctgttatg	cggccattgt	ccgtcaggac	2880
attgttggag	ccgaaatccg	cgtgcacgag	gtgccggact	tcggggcagt	cctcggccca	2940
aagcatcagc	tcacgcgag	cctgcgcgac	ggacgcactg	acggtgtcgt	ccatcacagt	3000
ttgccagtga	tacacatggg	gatcagcaat	cgcgcatatg	aatcacgcc	atgtagtgtg	3060
ttgaccgatt	ccttgccgtc	cgaatgggcc	gaaccgcctc	gtctggctaa	gatcggccgc	3120
agcgatcgca	tccatggcct	ccgcgaccgg	ctgcagaaca	gcgggcagtt	cggtttcagg	3180
caggtccttg	aacgtgacac	cctgtgcacg	gcgggagatg	caatagggtc	ggctctcgct	3240
aaattcccca	atgtcaagca	cttcgggaat	cgggagcgcg	gccgatgcaa	agtgccgata	3300
aacataacga	tctttgtaga	aaccatcggc	gcagctatct	accgcagga	catatccacg	3360
ccctcctaca	tcgaagctga	aagcacgaga	ttcttcgccc	tccgagagct	gcacaggtc	3420
ggagacgctg	tcgaactttt	cgatcagaaa	cttctcgaca	gacgtcgcg	tgagttcagg	3480
ctttttcatg	gtattatcat	cgtgtttttc	aaaggaaaac	cacgtccccg	tggttcgggg	3540
ggcctagacg	ttttttaacc	tcgactaaac	acatgtaaag	catgtgcacc	gaggccccag	3600
atcagatccc	atacaatggg	gtaccttctg	ggcatccttc	agccccttgt	tgaatacgct	3660
tgaggagagc	catttgactc	tttccacaac	tatccaactc	acaacgtggc	actgggggtg	3720
tgccgccttt	gcaggtgtat	cttatacacg	tggcttttgg	ccgcagaggc	acctgtcgcc	3780
aggtgggggg	ttccgctgcc	tgcaaagggt	cgctacagac	gttgtttgtc	ttcaagaagc	3840
ttccagagga	actgcttcc	tcacgacatt	caacagacct	tgcatctcct	tggcgagagg	3900
ggaaagaccc	ctagactaga	ccaagctttg	gatttcattt	ctgaagtttg	aattttctga	3960
gtcactagta	atgtccttga	ggatgatagt	ctgaattttc	tctgcaagag	tacaaagatt	4020
ggcttttttg	agatctttta	tcaatgtgtc	atacgttctc	ttctttccat	gaagttgatg	4080
ccaattacga	agcagttgaa	ctttctgttc	tgctgtgtct	tggacattgt	cattcttgat	4140
ctcatctatt	ttggcttcat	tgacaccatt	ctttcgaaca	aagcctttta	cttgacttag	4200
tgtcatgact	ccagcaatag	tggatgata	tttactcaag	tcaacatcag	ataaatttat	4260
tgccactgtt	tcaggattta	aggttgagaa	ttcatgagaa	ccttggtttt	cctttctgtg	4320
ctttctgcat	gttttctgta	cttcctttct	cttcacccaa	acaattagt	gaattggcaa	4380
aagaagaaga	caaagccacc	ccaaccggtt	tccggtcccc	ttcactgagc	cacggggccg	4440
acaatcttct	ggtctctggg	gctgagatgt	cccggtaggg	tgcacagggt	agggagttcg	4500
cagcactggc	ttggtagtag	tagagttcac	tttctgaagg	actggcacga	cagaactgaa	4560
gtacatcacc	gagttgctga	tgactgagca	gaaatagtag	ccttcgtttt	ccttgctgaa	4620
cttgttcagg	gtgagaacgt	acttattatt	cgtgtccctc	atggcagaaa	acagtttcga	4680
cgaattcagc	ttctcgctcc	acgttatctt	gttggtggat	gaagccatat	agacaacgaa	4740
ggtgggctgg	gggagtttgg	agctggagtt	ctggaagagc	caagagcatc	cttgcgaaac	4800
ggaccccaac	acttcacata	ccaggtccac	cttctgacca	agttcggcgt	ccattttctt	4860
tggaaagatt	cggagttcgg	gtgcctgtgg	cttagcttct	ccactcccca	ggataatcga	4920
ctcaccacgc	agcagcaggt	tcagcgacag	aaagcgggtc	aacggtgagg	ccatggtggc	4980
tttaccaaca	gtaccggaat	gccaaagctt	cggccgctta	agagctgtaa	ttgaacctgg	5040

gagtggacac	ctgtggagag	aaaggcaaag	tggatgtcag	taagaccaat	aggtgcctat	5100
cagaaacgca	agagtcttct	ctgtctcgac	aagcccagtt	tctatttggtc	tccttaaacc	5160
tgtcttgtaa	ccttgatact	tacctgcccc	gtgcctcacg	accaacttct	gcaggaattc	5220
ctggacagct	cccagatgat	cagtaaccgt	ggttgttatt	tctgtgccgg	gcagtggagc	5280
ctgggtaggg	ggagctctgc	ctcagtgcct	tcagctaaaa	atgggggtggg	aacccccagg	5340
aggcccgggc	cgccctggaa	gttccctttt	ctctctgttc	ttgggaagtc	gattgagcaa	5400
cagcgggggt	caggtgaggc	tccttcaacta	ccgatgcaca	ccgagtgcctg	ggggaggttc	5460
tcttctctct	caggcccaac	cccagggccc	ctgcctaggt	cccggactct	cactcttgac	5520
gcatgcgtgg	cttggtggtc	ccagtcagca	aacttggggg	cccgttgctt	gggaaaggga	5580
gagggtagctg	ggcatcgacg	cctctgcttc	cacgaaagcc	ttgtgaagaa	aggatggggg	5640
cgcttttgtg	caggagaatg	aggcgactg	aggatgaactg	gccctcgggg	gcgcgtgtcc	5700
cagatgtgtg	tgcagggcct	cctgatggcc	gcagccctcg	tcctgtgac	ccgcttgag	5760
ctggcaccct	gagtggtagc	ctcaccttgt	actcactccc	aggtcactgt	cctcgacgcg	5820
gccgctcgac	gataaaataa	aagattttat	ttagtctcca	gaaaaagggg	ggaatgaaag	5880
acccacactg	taggtttggc	aagctagctt	aagtaacca	ttttgcaagg	catgaaaaaa	5940
tacataactg	agaatagaga	agttcagatc	aaggtcgga	cagatggaac	aggcaataaa	6000
agagcccaca	acccctcact	cggggcgcca	gtcctccgat	tgactgagtc	gcccgggtac	6060
ccgtgtatcc	aataaacctt	cttgagttg	catccgactt	gtggtctcgc	tgttccttgg	6120
gagggctctcc	tctgagtgat	tgactaccgc	tcagcggggg	tctttcacat	gcagcatgta	6180
tcaaaattaa	tttggttttt	tttcttaagt	atttacatta	aatggccata	gtttcgtaat	6240
catggtcata	gctgtttcct	gtgtgaaatt	gttatccgct	cacaattcca	cacaacatac	6300
gagccggaag	cataaagtgt	aaagcctggg	gtgcctaattg	agtgagctaa	ctcacattaa	6360
ttgcgttgcg	ctcactgccc	gctttccagt	cgggaaacct	gtcgtgccag	ctgcattaat	6420
gaatcggcc	acgcgcgggg	agaggcggtt	tgcgatttgg	gcgctcttcc	gcttctcgc	6480
tactgactc	gctgcgctcg	gtcggtcggc	tgcggcgagc	ggtatcagct	cactcaaagg	6540
cggtaatagc	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaag	6600
gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	6660
gccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aaccgcagag	6720
gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	6780
ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcggtg	gcgctttctc	6840
atagctcacg	ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	6900
tgacgaacc	ccccgttcag	cccgaccgct	gcgccttctc	cggttaactat	cgtcttgagt	6960
ccaacccggt	aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	7020
gagcgaggta	tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	7080
ctagaaggac	agtattttgt	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	7140
ttggtagctc	ttgatccggc	aaacaaacca	ccgctggtag	cggtggtttt	tttgtttgca	7200
agcagcagat	tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	7260
ggtctgacgc	tcagtggaac	gaaaactcac	gttaagggat	tttggatcatg	agattatcaa	7320
aaaggatctt	cacctagatc	cttttaaatt	aaaaatgaag	tttgcgcaaa	tcaatctaaa	7380
gtatatatga	gtaaacttgg	tctgacagtt	accaatgctt	aatcagttag	gcacctatct	7440
cagcgatctg	tctatttctg	tcatccatag	ttgcctgact	ccccgtcgtg	tagataacta	7500
cgatacggga	gggcttacca	tctggcccc	gtgctgcaat	gataccgcga	gaccacgct	7560
caccggctcc	agatttatca	gcaataaacc	agccagccgg	aagggccgag	cgcagaagtg	7620
gtcctgcaac	tttatccgcc	tccatccagt	ctattaattg	ttgccgggaa	gctagagtaa	7680
gtagttcgcc	agttaatagt	ttgcgcaacg	ttgttgccat	tgctacaggc	atcgtggtgt	7740
cacgctcgtc	gtttggtatg	gcttcattca	gctccggttc	ccaacgatca	aggcgagtta	7800
catgatcccc	catgttgtgc	aaaaaagcgg	ttagctcctt	cggtcctccg	atcgttgtca	7860
gaagtaagtt	ggccgcagtg	ttatcactca	tggttatggc	agcactgcat	aattctctta	7920

ctgtcatgcc atccgtaaga tgcttttctg tgactgggtga gtactcaacc aagtcattct 7980  
 gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaacacgg gataataaccg 8040  
 cgccacatag cagaacttta aaagtgtca tcattggaaa acgttcttcg gggcgaaaac 8100  
 tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcaccaact 8160  
 gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 8220  
 atgccgcaaa aaaggaata agggcgacac ggaaatgttg aataactcata ctcttccttt 8280  
 ttcaatatta ttgaagcatt tatcagggtt attgtctcat gacattaacc tataaaaaata 8340  
 ggcgt 8345

<210> 9

<211> 61

<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: preferred  
 coiled-coil presentation structure

<400> 9

Met	Gly	Cys	Ala	Ala	Leu	Glu	Ser	Glu	Val	Ser	Ala	Leu	Glu	Ser	Glu
1					5				10				15		
Val	Ala	Ser	Leu	Glu	Ser	Glu	Val	Ala	Ala	Leu	Gly	Arg	Gly	Asp	Met
			20					25					30		
Pro	Leu	Ala	Ala	Val	Lys	Ser	Lys	Leu	Ser	Ala	Val	Lys	Ser	Lys	Leu
			35				40					45			
Ala	Ser	Val	Lys	Ser	Lys	Leu	Ala	Ala	Cys	Gly	Pro	Pro			
		50				55					60				

<210> 10

<211> 6

<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: loop structure  
 of coiled-coil presentation

<400> 10

Gly	Arg	Gly	Asp	Met	Pro
1				5	

<210> 11

<211> 69  
<212> PRT  
<213> Unknown Organism

<220>

<223> Description of Unknown Organism: preferred  
minibody presentation structure

<400> 11

Met Gly Arg Asn Ser Gln Ala Thr Ser Gly Phe Thr Phe Ser His Phe  
1 5 10 15

Tyr Met Glu Trp Val Arg Gly Gly Glu Tyr Ile Ala Ala Ser Arg His  
20 25 30

Lys His Asn Lys Tyr Thr Thr Glu Tyr Ser Ala Ser Val Lys Gly Arg  
35 40 45

Tyr Ile Val Ser Arg Asp Thr Ser Gln Ser Ile Leu Tyr Leu Gln Lys  
50 55 60

Lys Lys Gly Pro Pro  
65

<210> 12

<211> 7

<212> PRT

<213> Simian virus 40

<400> 12

Pro Lys Lys Lys Arg Lys Val  
1 5

<210> 13

<211> 6

<212> PRT

<213> Homo sapiens

<400> 13

Ala Arg Arg Arg Arg Pro  
1 5

<210> 14

<211> 10

<212> PRT



<213> Mus musculus

<400> 14

Glu Glu Val Gln Arg Lys Arg Gln Lys Leu  
1 5 10

<210> 15

<211> 9

<212> PRT

<213> Mus musculus

<400> 15

Glu Glu Lys Arg Lys Arg Thr Tyr Glu  
1 5

<210> 16

<211> 20

<212> PRT

<213> Xenopus laevis

<400> 16

Ala Val Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gly Ala Lys Lys  
1 5 10 15

Lys Lys Leu Asp  
20

<210> 17

<211> 10

<212> PRT

<213> Unknown Organism

<220>

<223> Description of Unknown Organism: preferred  
stability sequence

<220>

<221> UNSURE

<222> (3)..(6)

<223> "Xaa" at positions 3-6 can be any amino acid.

<400> 17

Met Gly Xaa Xaa Xaa Xaa Gly Gly Pro Pro  
1 5 10

<210> 18  
<211> 5  
<212> PRT  
<213> Unknown Organism

<220>  
<223> Description of Unknown Organism: linker consensus  
sequence

<400> 18  
Gly Ser Gly Gly Ser  
1 5

<210> 19  
<211> 4  
<212> PRT  
<213> Unknown Organism

<220>  
<223> Description of Unknown Organism: linker consensus  
sequence

<400> 19  
Gly Gly Gly Ser  
1